## **CLAIMS**

,	\A/ A	$\sim$	0110	•
<u>.</u>	W C	•	laim	

3	1.	An apparatus for switching packets, each packet having a header
4	portion, an optic	onal corresponding tail portion, and a class of service indicator, said
5	apparatus comp	rising:
6	a pipelir	ned switch comprising:
7	a	plurality of packet header buffers (PHBs);
8	a	un equal plurality of PHB pointers, each said PHB pointer pointing to a
9		corresponding PHB; and
10	a	in equal plurality of pipeline stage circuits connected in a sequence,
11		comprising at least a first stage circuit and a last stage circuit,
12	,	vherein:
13 <u>)</u>	/ X)	each said stage circuit begins an operation substantially
14_	<b>3</b> (\	simultaneously with each other;
15	0,	each said stage circuit passes data to a next stage circuit in said
16	4	sequence when every said operation performed by all
17		said stage circuits is completed;
18		said first stage circuit reads said header portion and stores said
19		header portion in said corresponding PHB using said
20		corresponding PHB pointer; and
21		said last stage circuit outputs a modified header portion; and
22	a receive	buffer manager (RBM) comprising:
23	а	joining circuit connected to said pipelined switch wherein said
24		modified header portion and said corresponding tail portion are
25		joined to form a modified packet;
26	а	receive queue manager connected to said joining circuit that buffers
27		said modified packet in a receive packet buffer and enqueues
28		said modified packet using said class of service indicator and a
29		plurality of receive queues; and

a de

5

6 7

8 9

10

11

12

13 14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

a dequeue circuit connected to said receive queue manager and said receive packet buffer, wherein said dequeue circuit uses said class of service indicator to dequeue said modified packet to a switch fabric.

- The apparatus as recited in Claim 1, wherein said plurality of pipeline 2. stage circuits further comprise: a pre-process circuit connected to said first stage circuit, wherein said preprocess circuit uses a second said PHB pointer to record first data in said corresponding PHB; a pointer lookup circuit connected to said pre-process circuit that compares said header portion to a first data structure and determines a leaf pointer using said second PHB pointer; a table lookup circuit connected to said pointer lookup circuit that uses said leaf pointer to access one or more sets of linked data structures and to fetch second data, wherein said table lookup circuit uses a third said PHB pointer to record said second data in said corresponding PHB; and a post-process circuit using said third PHB pointer and connected to said table lookup circuit, wherein said post-process circuit uses a fourth said PHB pointer to record third data in said corresponding PHB;
- wherein said last pipeline stage circuit comprises a gather circuit connected to said post-process circuit, and wherein said gather circuit uses said fourth PHB pointer to assemble said modified header portion.
  - 3. The apparatus as recited in Claim 1, further comprising: an input device that receives said packet; and a first buffer connected between said input device and said first stage circuit, wherein said first buffer buffers said header portion and said tail portion.

1	4.	The apparatus as recited in Claim 3, further comprising a multiplexer
2	interposed be	etween said first buffer and said first stage circuit, wherein said
3	multiplexer i	multiplexes said header portion and said corresponding tail portion
4	together.	
5	5.	The apparatus as recited in Claim 1, further comprising a second buffer
6	interposed be	etween said last stage circuit and said joining circuit.
7	6.	The apparatus as recited in Claim 5, further comprising a multiplexer
8	interposed be	etween said last stage circuit and said second buffer, wherein said
9	multiplexer i	nultiplexes said modified header portion and said corresponding tail
10	portion toget	her.
11	7.	The apparatus as recited in Claim 1, wherein said receive packet buffer
12	comprises bu	affers of different sizes.
	0	
13	8.	The apparatus as recited in Claim 1, wherein said receive packet buffer
14	comprises bu	affers of equal size.
15	9.	The apparatus as recited in Claim 1, wherein said receive queue
16		apprises a congestion avoidance circuit utilizing a status of each said
17	receive queu	·
.,	rocerve queu	
18	10.	The apparatus as recited in Claim 9, wherein said status comprises a
19	measure of a	verage queue depth.
20		
21	11.	The apparatus as recited in Claim 1, further comprising a transmit
22	buffer manag	ger (TBM), said TBM comprising:
23	a thir	d buffer that receives one or more packets from said switch fabric;
24	a tran	smit queue manager connected to said third buffer that buffers each said
25		packet in a transmit packet buffer and enqueues said packet using said
26		class of service indicator and a plurality of transmit queues; and

I	a deq	ueue circuit connected to said transmit queue manager and said transmit
2 ·		packet buffer, wherein said dequeue circuit uses said class of service
3		indicator to dequeue each said packet.
4	12.	The apparatus as recited in Claim 11, wherein said transmit packet
5	buffer compr	ises buffers of different sizes.
6	13.	The apparatus as recited in Claim 11, wherein said transmit packet
7	buffer compr	ises buffers of equal size.
8	14.	The apparatus as recited in Claim 11, wherein said transmit queue
9	manager com	prises a congestion avoidance circuit utilizing a status of each said
10	transmit queu	ne.
11	15.	The apparatus as recited in Claim 14, wherein said status comprises a
12	measure of av	verage queue depth.
13	16.	The apparatus as recited in Claim 11, further comprising a transmit
14	FIFO connec	ted to an output of said dequeue circuit.
15		
16		
17	17.	An apparatus for switching packets in a communications network
18	device compr	rising:
19	a buff	er that receives one or more packets, said packets comprising a class of
20		service indicator;
21	a trans	smit queue manager connected to said buffer that buffers each said
22		packet in a transmit packet buffer and enqueues said packet using said
23		class of service indicator and a plurality of transmit queues; and
24	a dequ	ueue circuit connected to said transmit queue manager and said transmit
25		packet buffer, wherein said dequeue circuit uses said class of service
26		indicator to dequeue each said packet.

		1
1	18.	The apparatus as recited in Claim 17, wherein said transmit packet
2	buffer compri	ises buffers of different sizes.
3	19.	The apparatus as recited in Claim 17, wherein said transmit packet
4		ises buffers of equal size.
7	ourier compri	ses buriers of equal size.
5	20.	The apparatus as recited in Claim 17, wherein said transmit queue
6	manager com	prises a congestion avoidance circuit utilizing a status of each said
7	transmit queu	e.\
8	21.	The apparatus as recited in Claim 20, wherein said status comprises a
9	measure of av	rerage queue depth.
10	22.	The apparatus as recited in Claim 1/7, further comprising a transmit
11	FIFO connect	ed to an output of said dequeue circuit.
12		
13		\
	/	
14	23.	A method of switching packets, which comprises:
15	receivi	ing a packet, said packet having a header portion, an optional
16		corresponding tail portion, and a class of service indicator;
17	switch	ing said packet through a pipelined switch having a plurality of packet
18	1/	header buffers (PHBs), an equal plurality of PHB pointers wherein
12/		each said PHB pointer points to a corresponding PHB, and an equal
ر 20	\ \ \	plurality of pipeline stages connected in a sequence, comprising at least
21	$\bigcirc$	a first stage and a last stage, said switching further comprising:
22	K, /	beginning said sequence in each said stage substantially simultaneously
23		with each other said stage;
24	V.	passing data to a next stage circuit in said sequence when every said
25		operation performed by all said stage circuits is completed;
26	$\bigcirc$	reading and storing said header in said corresponding PHB using said
27		corresponding PHB pointer; and
28		outputting a modified header portion; and
		- 79 -

21

1	buffering said modified header portion in a receive buffer manager (RBM),
2	said buffering further comprising:
3	joining said modified header portion and said corresponding tail
4	portion to form a modified packet;
5	buffering and enqueuing said modified packet using said class of
6	service indicator; and
7	dequeuing said modified packet using said class of service indicator.
8	24. The method of Claim 23, wherein said switching further comprises:
9	recording first data in said corresponding PHB using a second said PHB
10	pointer;
11	comparing said header portion to a first data structure and determining a leaf
12	pointer using said second PHB pointer;
13	fetching second data using said leaf pointer to access one or more sets of
14	linked data structures and recording said second data in said
15	corresponding PHB using a third said PHB pointer;
16	post-processing said header portion using said third PHB pointer and recording
17	third data in said corresponding PHB using a fourth said PHB pointer;
18	and
19	assembling said modified header portion using said fourth PHB pointer.
20	25. The method of Claim 23, wherein said switching further comprises

The method of Claim 25, wherein said switching further comprises 26. multiplexing said header portion and said tail portion together prior to beginning said sequence.

buffering said header portion and said tail portion prior to said switching.

- 25 27. The method of Claim 23, wherein said buffering further comprises 26 buffering said header portion and said tail portion prior to said joining.
- 27 28. The method of Claim 27, wherein said buffering further comprises 28 multiplexing said header portion and said tail portion together prior to said joining.

1	29. The method of Claim 23, further comprising:
2	receiving one or more packets;
3	buffering and enqueuing each said packet using said class of service indicator
4	and a plurality of queues; and
5	dequeuing each said packet using said class of service indicator.
6	30. The method of Claim 29, wherein said buffering uses a packet buffer
7	comprising buffers of different sizes.
8	31. The method of Claim 29, wherein said buffering uses a packet buffer
9	comprising buffers of equal size.
10	32. The method of Claim 29, wherein said buffering further comprises
11	avoiding congestion using a status of each said queue.
12	33. The method of Claim 32, wherein said status comprises a measure of
13	average queue depth.
14	34. The method of Claim 29, wherein said dequeuing uses a transmit FIFO.
15	
16	
17	35 A method of avitching pockets, which comprises:
18	35. A method of switching packets, which comprises: receiving one or more packets having a class of service indicator;
19	buffering and enqueuing each said packet using said class of service indicator
20	and a plurality of queues; and
21	dequeuing each said packet using said class of service indicator.
<b>4</b> 1	dequeding each said packet using said class of service materior.
22	36. The method of Claim 35, wherein said buffering uses a packet buffer
23	comprising buffers of different sizes.
24	37. The method of Claim 35, wherein said buffering uses a packet buffer
2 <del>4</del> 25	comprising buffers of equal size.
	comprising contain of odom hine.

1	38.	The method of Claim 35, wherein said buffering further comprises
2	avoiding con	gestion using a status of each said queue.
3	39.	The method of Claim 38, wherein said status comprises a measure of
4	average queu	e depth.
5	40.	The method of Claim 35, wherein said dequeuing uses a transmit FIFO
6		$\sum_{i}$
7		
8	41.	A computer system for interfacing with a communications network,
9	comprising co	omputer instructions for:
10	receiv	ring a packet, said packet having a header portion, an optional
11		corresponding tail portion, and a class of service indicator;
12	switch	ning said packet through a pipelined switch having a plurality of packet
13 \		header buffers (PHBs), an equal plurality of PHB pointers wherein
14	$\bigvee$	each said PHB pointer points to a corresponding PHB, and an equal
13/		plurality of pipeline stages connected in a sequence, comprising at least
16		a first stage and a last stage, said switching further comprising:
17	V /	beginning said sequence in each said stage substantially simultaneously
18	<b>K</b> ,	with each other said stage;
19	·	passing data to a next stage circuit in said sequence when every said
20		operation performed by all said stage circuits is completed;
21		reading and storing said header in said corresponding PHB using said
22		corresponding PHB pointer; and
23		outputting a modified header portion; and
24	buffer	ing said modified header portion in a receive buffer manager (RBM),
25		said buffering further comprising:
26		joining said modified header portion and said corresponding tail
27		portion to form a modified packet;
28		buffering and enqueuing said modified packet using said class of
29		service indicator; and
		1

1	dequeuing said modified packet using said class of service indicator.
2	42. The computer system of Claim 41, wherein said switching further
3	comprises:
4	recording first data in said corresponding PHB using a second said PHB
5	pointer;
6	comparing said header portion to a first data structure and determining a leaf
7	pointer using said second PHB pointer;
8	fetching second data using said leaf pointer to access one or more sets of
9	linked data structures and recording said second data in said
10	corresponding PHB using a third said PHB pointer;
11	post-processing said header portion using said third PHB pointer and recording
12	third data in said corresponding PHB using a fourth said PHB pointer;
13	and
14	assembling said modified header portion using said fourth PHB pointer.
15	43. The computer system of Claim 41, further comprising:
16	receiving one or more packets;
17	buffering and enqueuing each said packet using said class of service indicator
18	and a plurality of queues; and
19	dequeuing each said packet using said class of service indicator.
20	
21	
	\
22	44. A computer readable storage medium, comprising computer

) instructions for:

27

28

receiving a packet, said packet having a header portion, an optional corresponding tail portion, and a class of service indicator; switching said packet through a pipelined switch having a plurality of packet header buffers (RHBs), an equal plurality of PHB pointers wherein

each said PHB pointer points to a corresponding PHB, and an equal

1	plurality of pipeline stages connected in a sequence, comprising at least
2	a first stage and a last stage, said switching further comprising:
3	beginning said sequence in each said stage substantially simultaneously
4	with each other said stage;
5	passing data to a next stage circuit in said sequence when every said
6	operation performed by all said stage circuits is completed;
	reading and storing said header in said corresponding PHB using said
8)	corresponding PHB pointer; and
9	outputting a modified header portion; and
10	buffering said modified header portion in a receive buffer manager (RBM),
11	said buffering further comprising:
12	joining said modified header portion and said corresponding tail
13	portion to form a modified packet;
14	buffering and enqueuing said modified packet using said class of
15	service indicator; and
16	dequeuing said modified packet using said class of service indicator.
17	45. The computer readable storage medium of Claim 44, wherein said
18	switching further comprises:
19	recording first data in said corresponding PHB using a second said PHB
20	pointer;
21	comparing said header portion to a first data structure and determining a leaf
22	pointer using said second PHB pointer;
23	fetching second data using said leaf pointer to access one or more sets of
24	linked data structures and recording said second data in said
25	corresponding PHB using a third said PHB pointer;
26	post-processing said header portion using said third PHB pointer and recording
27	third data in said corresponding PHB using a fourth said PHB pointer;
28	and
29	assembling said modified header portion using said fourth PHB pointer.

1	4	5. The computer readable storage medium of Claim 44, further
2	comprisi	ng:
3	re	ceiving one or more packets;
4	b	uffering and enqueuing each said packet using said class of service indicator
5		and a plurality of queues; and
6	de	equeuing each said packet using said class of service indicator.
7		
8		
9	4′	A computer data signal embodied in a carrier wave, comprising
10	computer	instructions for:
11	re	ceiving a packet, said packet having a header portion, an optional
12		corresponding tail portion, and a class of service indicator;
13	sv	vitching said packet through a pipelined switch having a plurality of packet
14		header buffers (PHBs), an equal plurality of PHB pointers wherein
15 \	\	each said PHB pointer points to a corresponding PHB, and an equal
16	$\mathcal{N}$	plurality of pipeline stages connected in a sequence, comprising at leas
17		a first stage and a last stage, said switching further comprising:
18	W/	beginning said sequence in each said stage substantially simultaneously
19	1,4	with each other said stage;
20	UN.	passing data to a next stage circuit in said sequence when every said
21	(10)	operation performed by all said stage circuits is completed;
22	$\circ$	reading and storing said header in said corresponding PHB using said
23		corresponding PHB pointer; and
24		outputting a modified header portion; and
25	bı	offering said modified header portion in a receive buffer manager (RBM),
26		said buffering further comprising:
27		joining said modified header portion and said corresponding tail
28		portion to form a modified packet;
29		buffering and enqueuing said modified packet using said class of
30		service indicator; and
31		dequeuing said modified packet using said class of service indicator.

1	48.	The computer data signal of Claim 47, wherein said switching further
2	comprises:	
3	record	ling first data in said corresponding PHB using a second said PHB
4		pointer;
5	compa	aring said header portion to a first data structure and determining a leaf
6		pointer using said second PHB pointer;
7	fetchi	ng second data using said leaf pointer to access one or more sets of
8		linked data structures and recording said second data in said
9		corresponding PHB using a third said PHB pointer;
10	post-p	processing said header portion using said third PHB pointer and recording
11		third data in said corresponding PHB using a fourth said PHB pointer;
12		and
13	assem	bling said modified header portion using said fourth PHB pointer.
14	49.	The computer data signal of Claim 47, further comprising:
15		ing one or more packets;
16		ing and enqueuing each said packet using said class of service indicator
17	builei	and a plurality of queues; and
18	deane	uing each said packet using said class of service indicator.
19	acque	unig each said packet using said class of service indicator.
20		
20		
21	50.	An apparatus for switching packets, said packets having a header
22	portion, a tail	portion, and a class of service indicator, comprising:
23		lined switch comprising a plurality of stage circuits connected in a
24	• •	sequence wherein:
25		each said stage circuit begins an operation substantially simultaneously
26		with each other said stage circuit;
27		each said stage circuit passes data when every said operation performed
28		by all said stage circuits is completed;
29	a last :	stage circuit that outputs a modified header portion;
		, '\ '\

29

l	a queue manager connected to said last stage circuit, wherein said queue
2	manager
3	joins said modified header portion and said tail portion to form a
1	modified packet; and
5	routes said modified packet using said class of service indicator.
	\